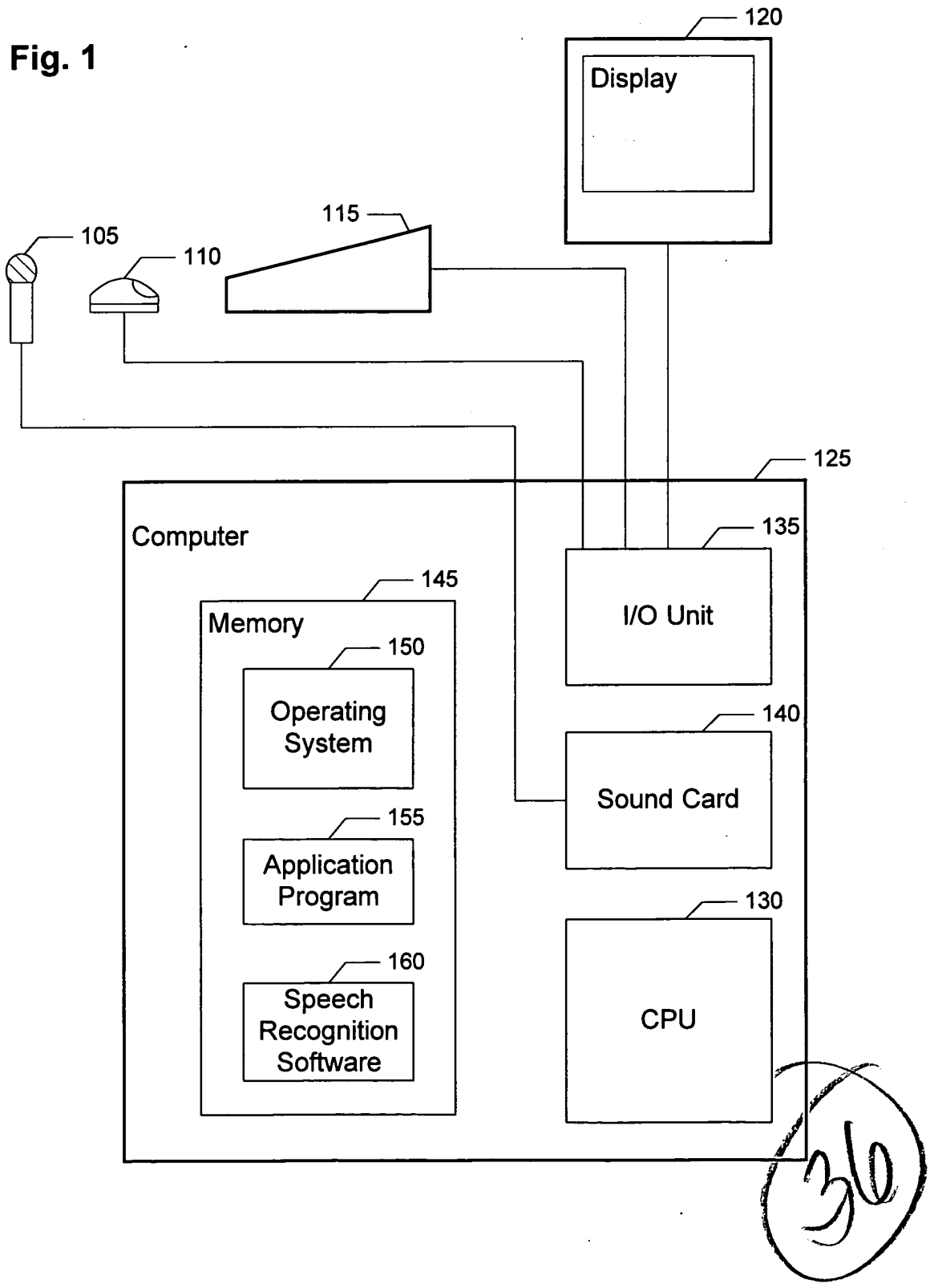
704/251

Fig. 1


120
Display
115
105
110
125
Computer
135
I/O Unit
145
Memory
150
Operating System
140
Sound Card
155
Application Program
130
160
Speech Recognition Software
CPU
30

EP 0 962 920 A2

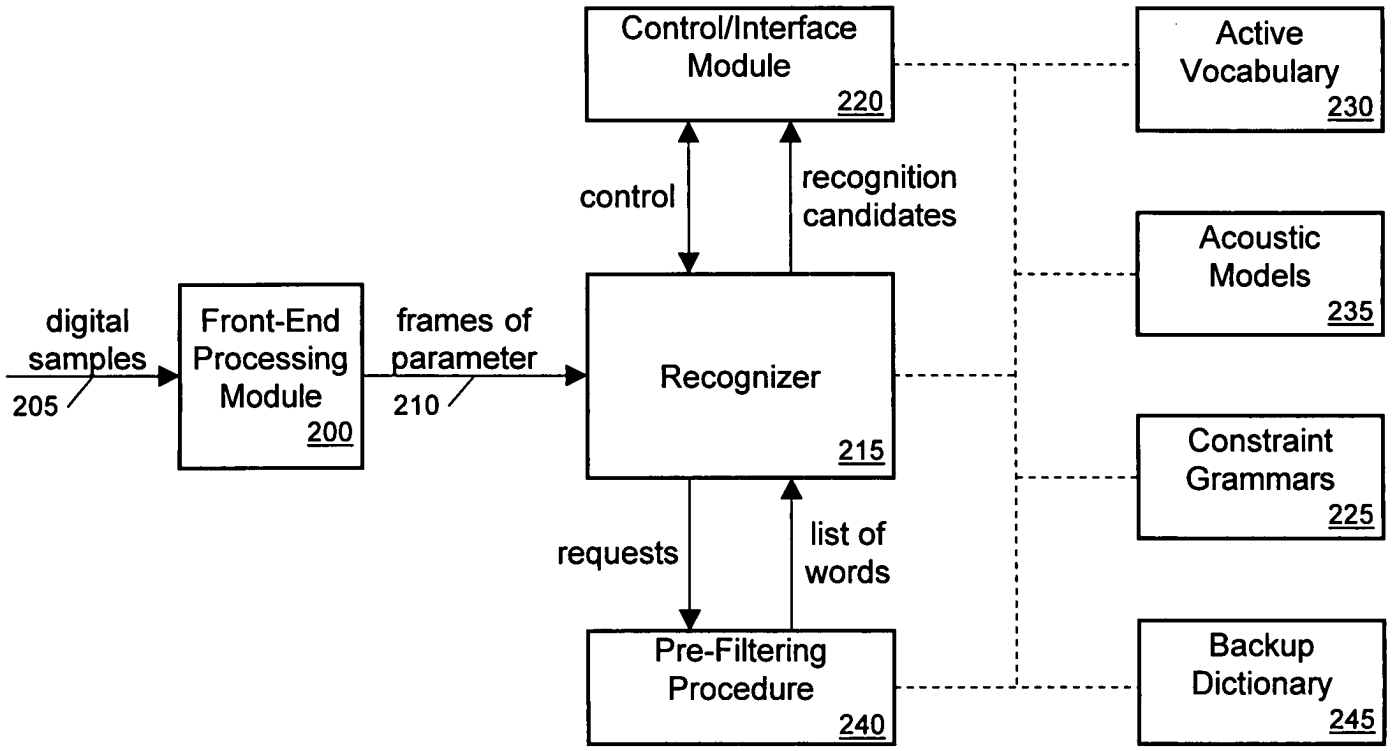


Fig. 2

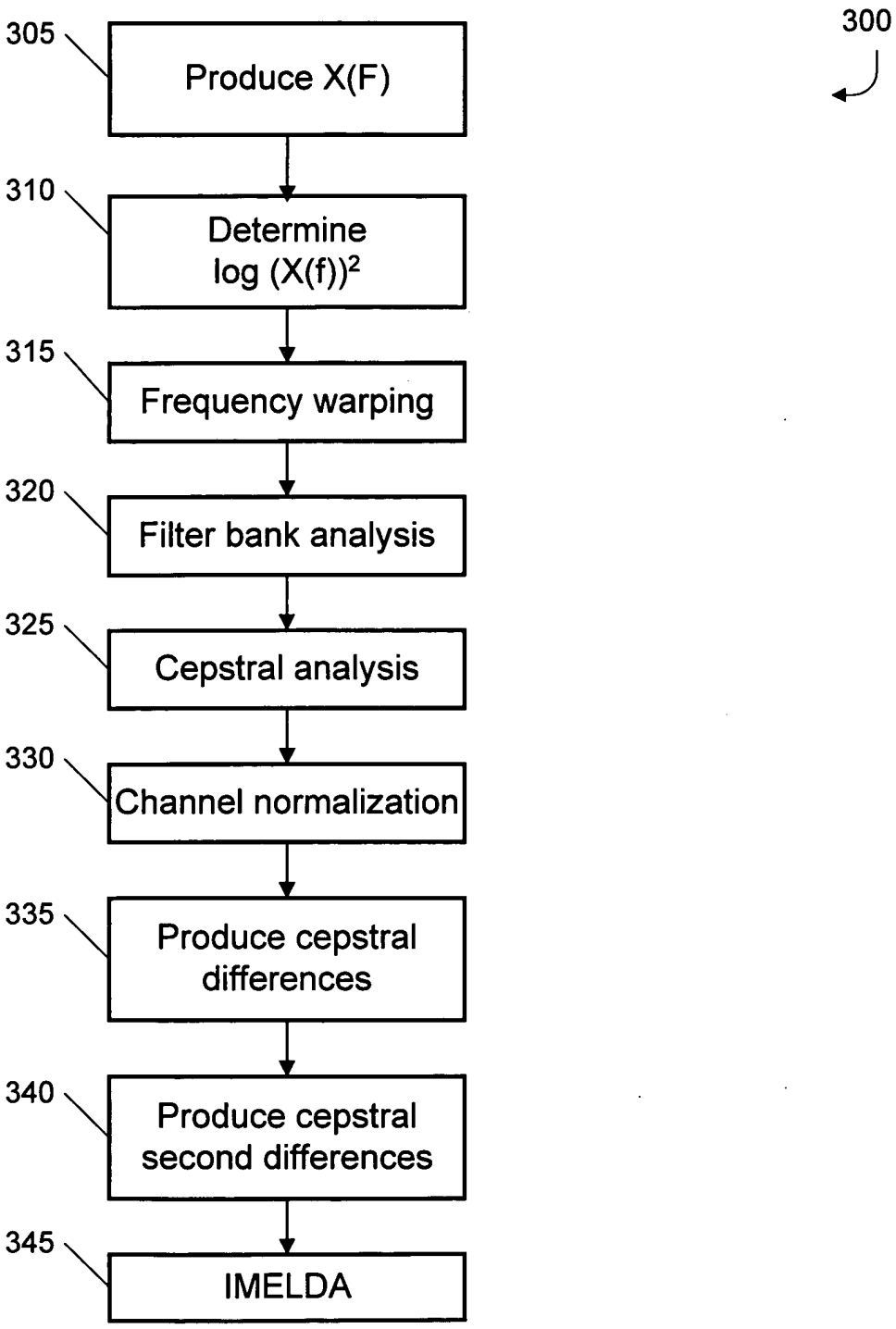

300
305
Produce X(F)
310
Determine
log (X(f))²
315
Frequency warping
320
Filter bank analysis
325
Cepstral analysis
330
Channel normalization
335
Produce cepstral
differences
340
Produce cepstral
second differences
345
IMELDA

Fig. 3

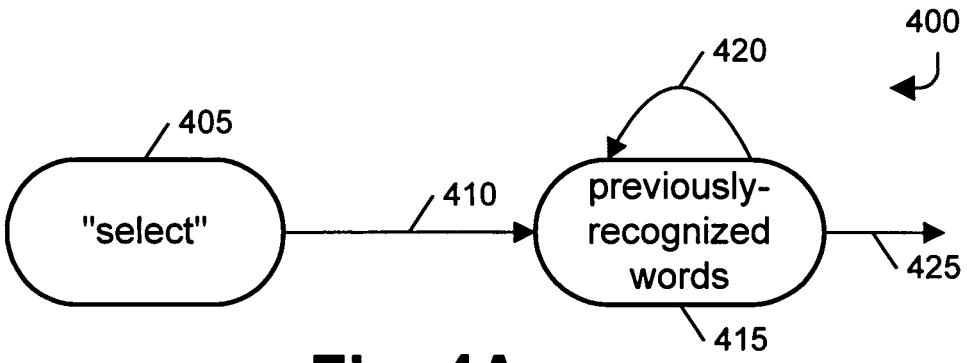

400
420
405
"select"
410
previously-recognized words
425
415

Fig. 4A

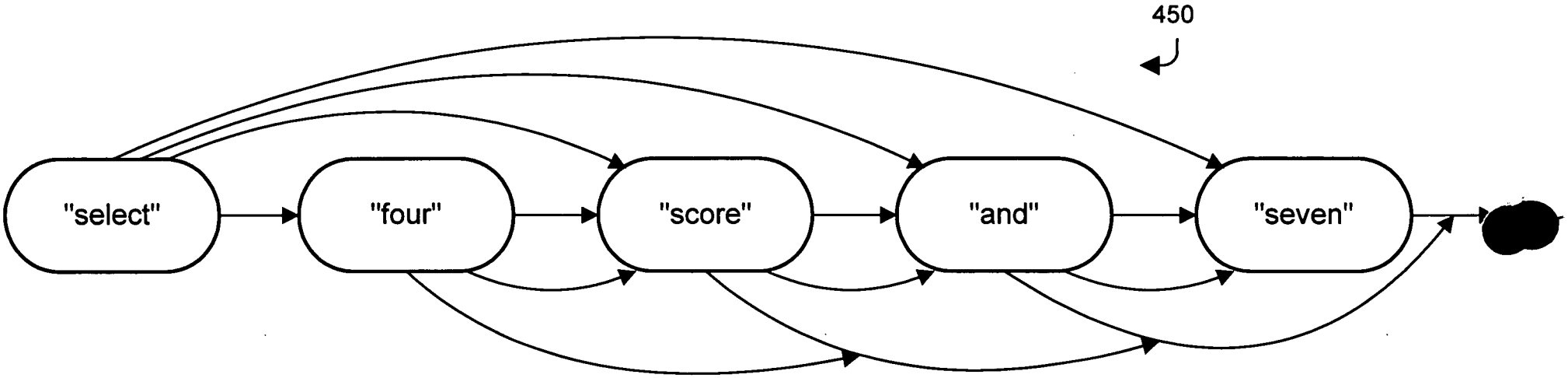

450
"select"
"four"
"score"
"and"
"seven"

Fig. 4B

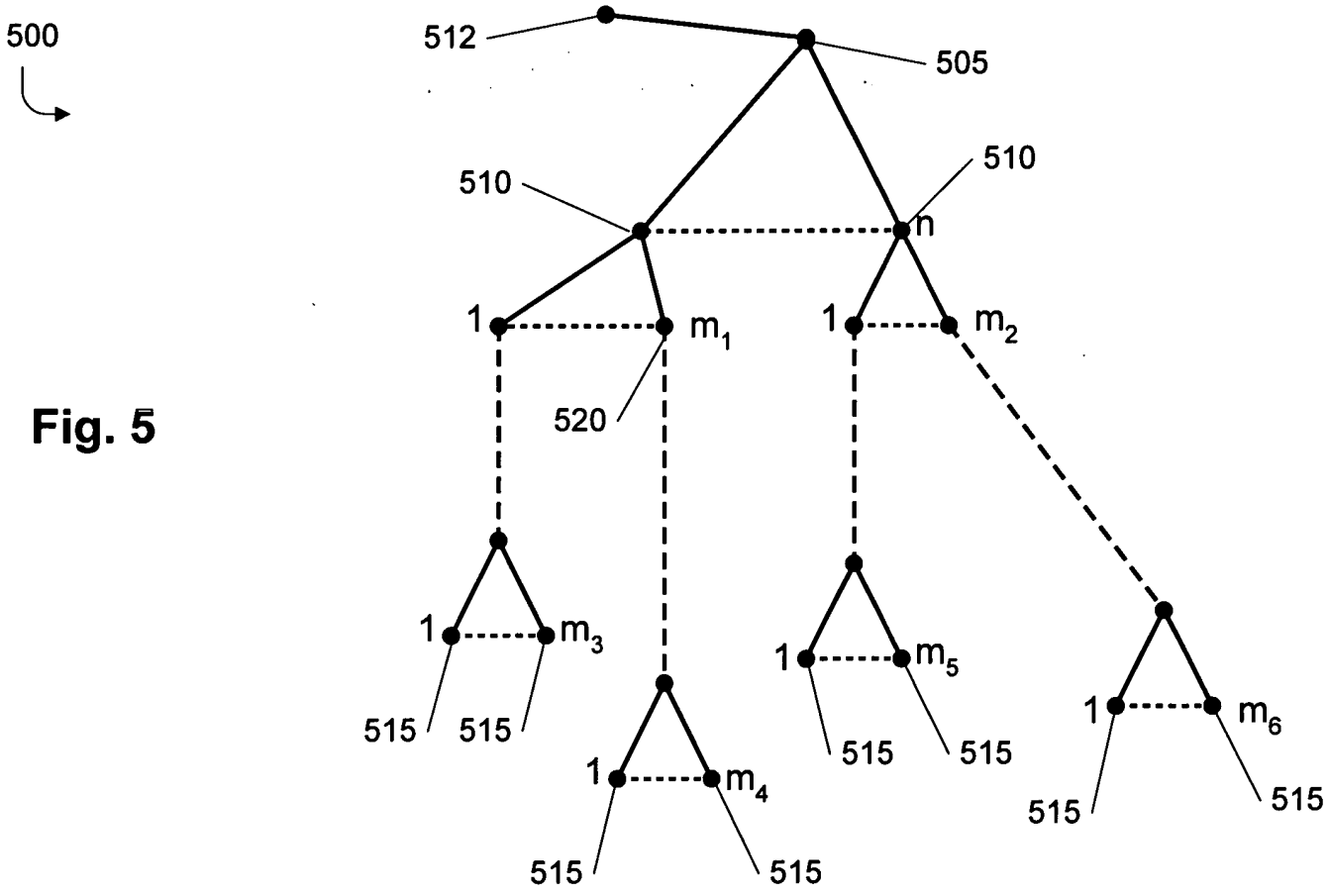

500
512
505
510
510
n
1
m1
1
m2
520
1
m3
1
m5
1
m6
515
515
1
m4
515
515
515
515
515

Fig. 5

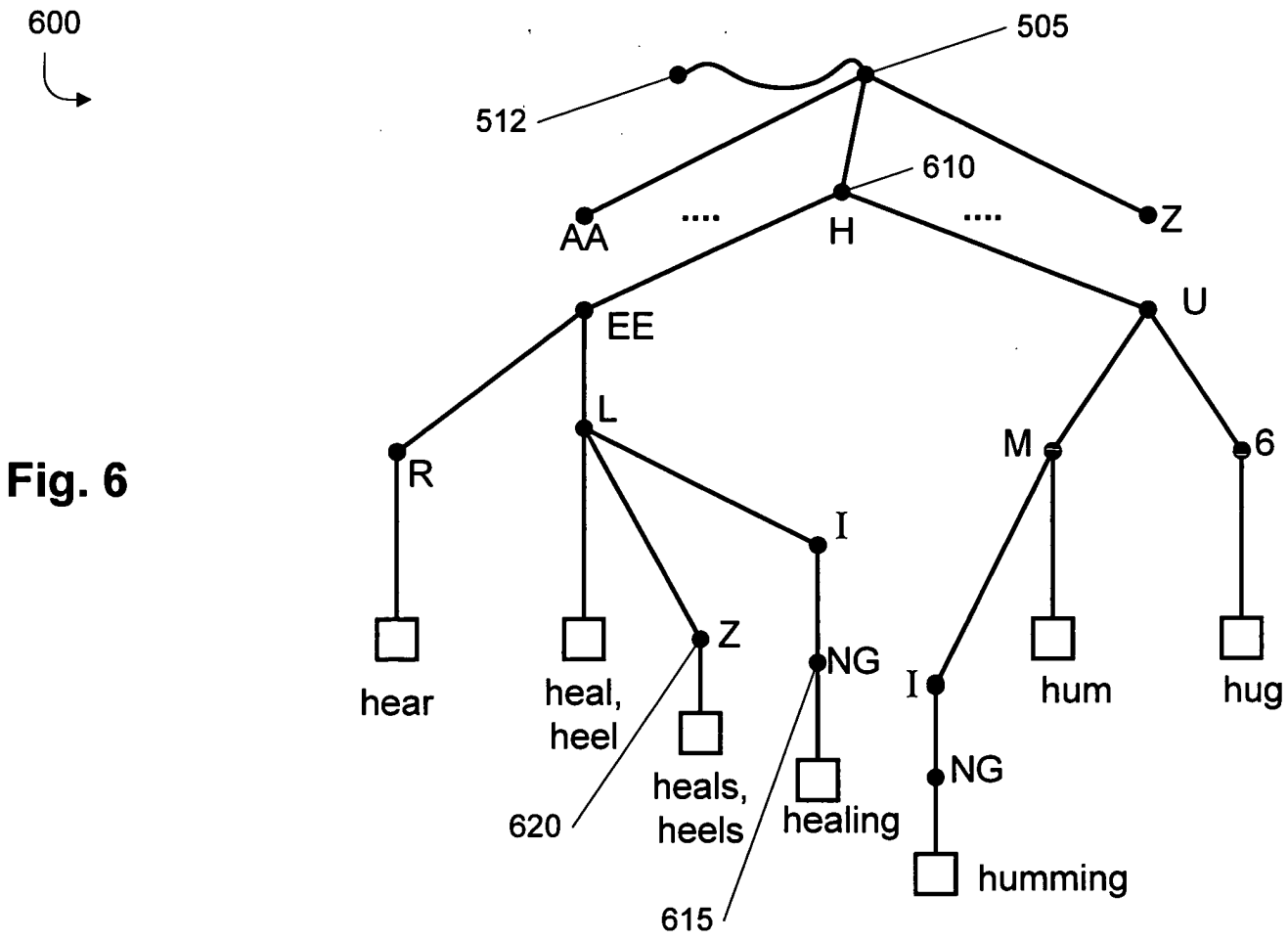

600
505
512
610
AA
....
H
....
Z
EE
U
L
R
M
6
I
Z
NG
I
hear
heal,
heel
hum
hug
620
heals,
heels
healing
NG
humming
615

Fig. 6

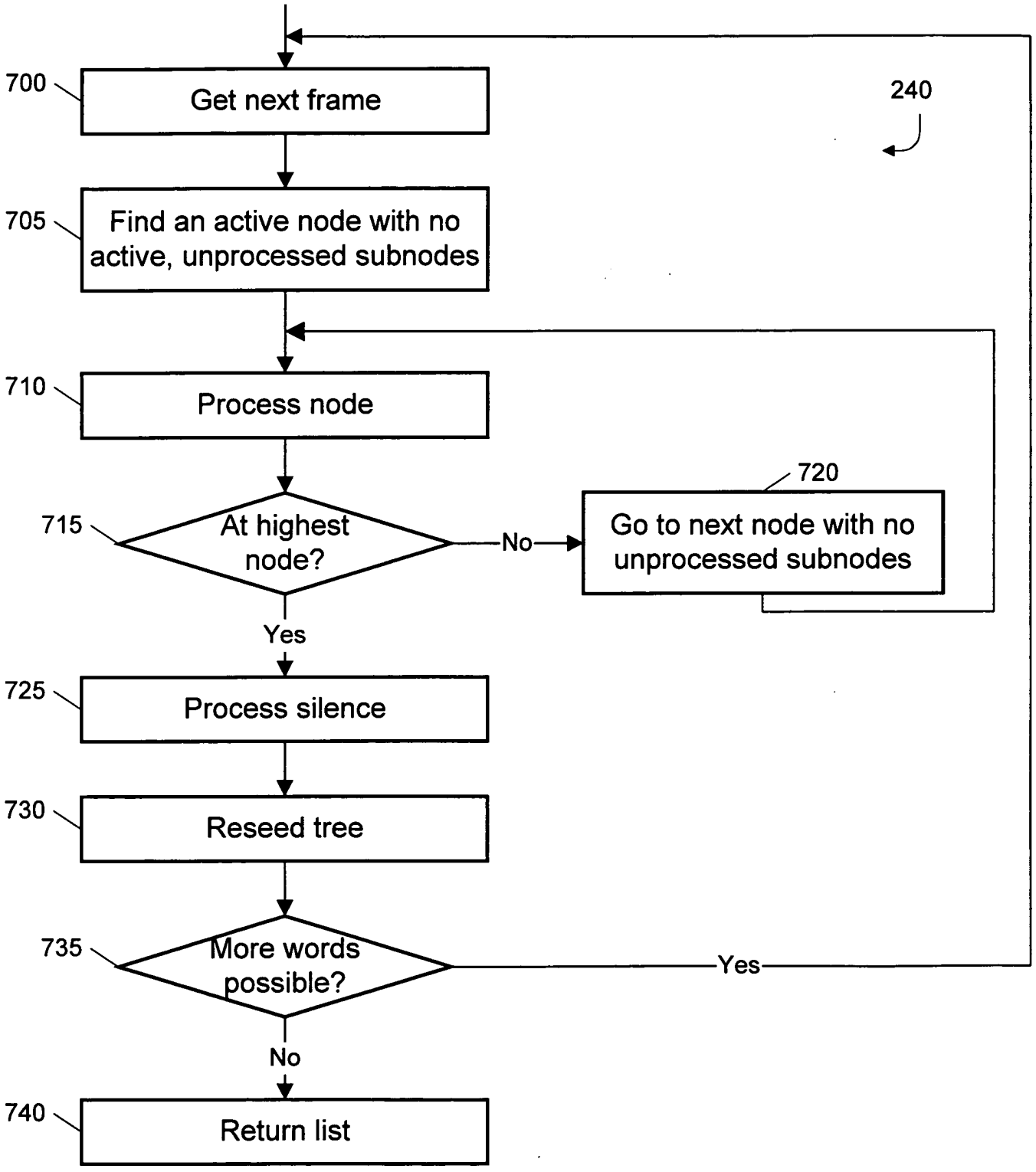

700
Get next frame
240
705
Find an active node with no active, unprocessed subnodes
710
Process node
720
715
At highest node?
No
Go to next node with no unprocessed subnodes
Yes
725
Process silence
730
Reseed tree
735
More words possible?
Yes
No
740
Return list

Fig. 7

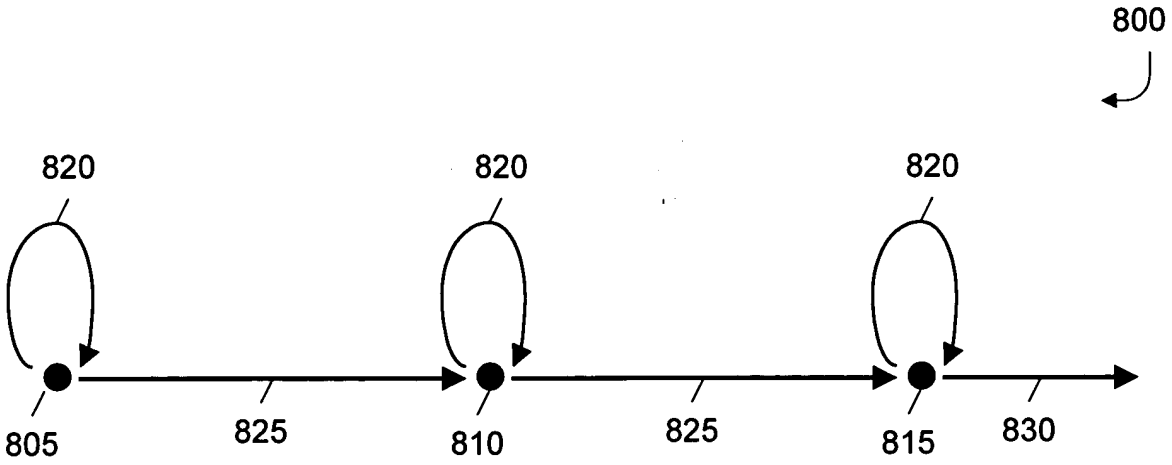


Fig. 8A

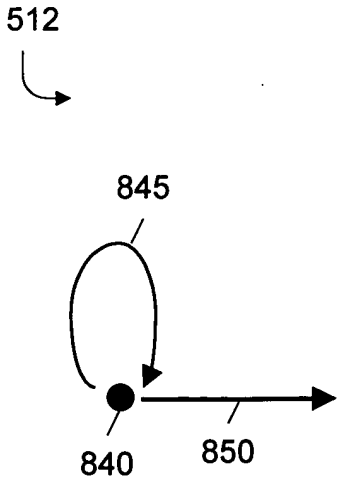


Fig. 8B

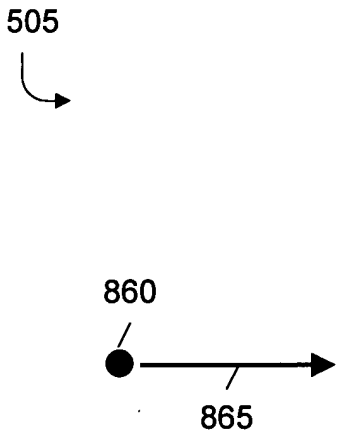


Fig. 8C

	Frame	840 ("A")	805 ("B")	810 ("C")	815 ("D")	Next Node ("N")
900	0	0	-----	-----	-----	-----
905	1	$S_{A1}=A_{A1}$	$S_{B1}=A_{B1}$	-----	-----	-----
910	2	$S_{A2}=S_{A1}+A_{A2}$	$S_{B2}=\min(S_{B1}+stay_B, S_{A1})+A_{B2}$	$S_{C2}=S_{B1}+leave_B+A_{C2}$	-----	-----
915	3	$S_{A3}=S_{A2}+A_{A3}$	$S_{B3}=\min(S_{B2}+stay_B, S_{A2})+A_{B3}$	$S_{C3}=\min(S_{C2}+stay_C, S_{B2}+leave_B)+A_{C3}$	$S_{D3}=S_{C2}+leave_C+A_{D3}$	-----
920	4	$S_{A4}=S_{A3}+A_{A4}$	$S_{B4}=\min(S_{B3}+stay_B, S_{A3})+A_{B4}$	$S_{C4}=\min(S_{C3}+stay_C, S_{B3}+leave_B)+A_{C4}$	$S_{D4}=\min(S_{D3}+stay_D, S_{c3}+leave_C)+A_{D4}$	$S_{N4}=S_{D2}+leave_D+A_{D4}$
	. . .					
925	n	$S_{An}=S_{An-1}+A_{An}$	$S_{Bn}=\min(S_{Bn-1}+stay_B, S_{An-1})+A_{Bn}$	$S_{Cn}=\min(S_{Cn-1}+stay_C, S_{Bn-1}+leave_B)+A_{Cn}$	$S_{Dn}=\min(S_{Dn-1}+stay_D, +S_{cn-1}+leave_C)+A_{Dn}$	$S_{NN}=\min(S_{Nn-1}+stay_M, S_{Mn-1}+leave_M)+A_{Nm}$

Fig. 9

	Frame	810 ("A")	805 ("B")	810 ("C")	815 ("D")	Next Node ("N")
900	0	$S_{AO}=0$	-----	-----	-----	-----
905	1	$S_{A1}=f(S_{AO}, A_{A1})$	$S_{B1}=f(S_{AO}, A_{B1})$	-----	-----	-----
910	2	$S_{A2}=f(S_{A1}, A_{A2})$	$S_{B2}=f(S_{B1}, stay_B, S_{A1}, A_{B2})$	$S_{C2}=f(S_{B1}, leave_B, A_{C2})$	-----	-----
915	3	$S_{A3}=f(S_{A2}, A_{A3})$	$S_{B3}=f(S_{B2}, stay_B, S_{A2}, A_{B3})$	$S_{C3}=f(S_{C2}, stay_C, S_{B2}, leave_B, A_{C3})$	$S_{D3}=f(S_{C2}, leave_C, A_{D3})$	-----
920	4	$S_{A4}=f(S_{A3}, A_{A4})$	$S_{B4}=f(S_{B3}, stay_B, S_{A3}, A_{B4})$	$S_{C4}=f(S_{C3}, stay_C, S_{B3}, leave_B, A_{C4})$	$S_{D4}=f(S_{D3}, stay_D, S_{c3}, leave_C, A_{D4})$	$S_{N4}=f(S_{D3}, leave_D, A_{D3})$
	. . .					
925	n	$S_{An}=f(S_{An-1}, A_{An})$	$S_{Bn}=f(S_{Bn-1}, stay_B, S_{An-1}, A_{Bn})$	$S_{Cn}=f(S_{Cn-1}, stay_C, S_{Bn-1}, leave_B, A_{Cn})$	$S_{Dn}=f(S_{Dn-1}, stay_D, S_{cn-1}, leave_C, A_{Dn})$	$S_{Nn}=f(S_{Dn-1}, leave_N, A_{Nn})$

Fig. 10

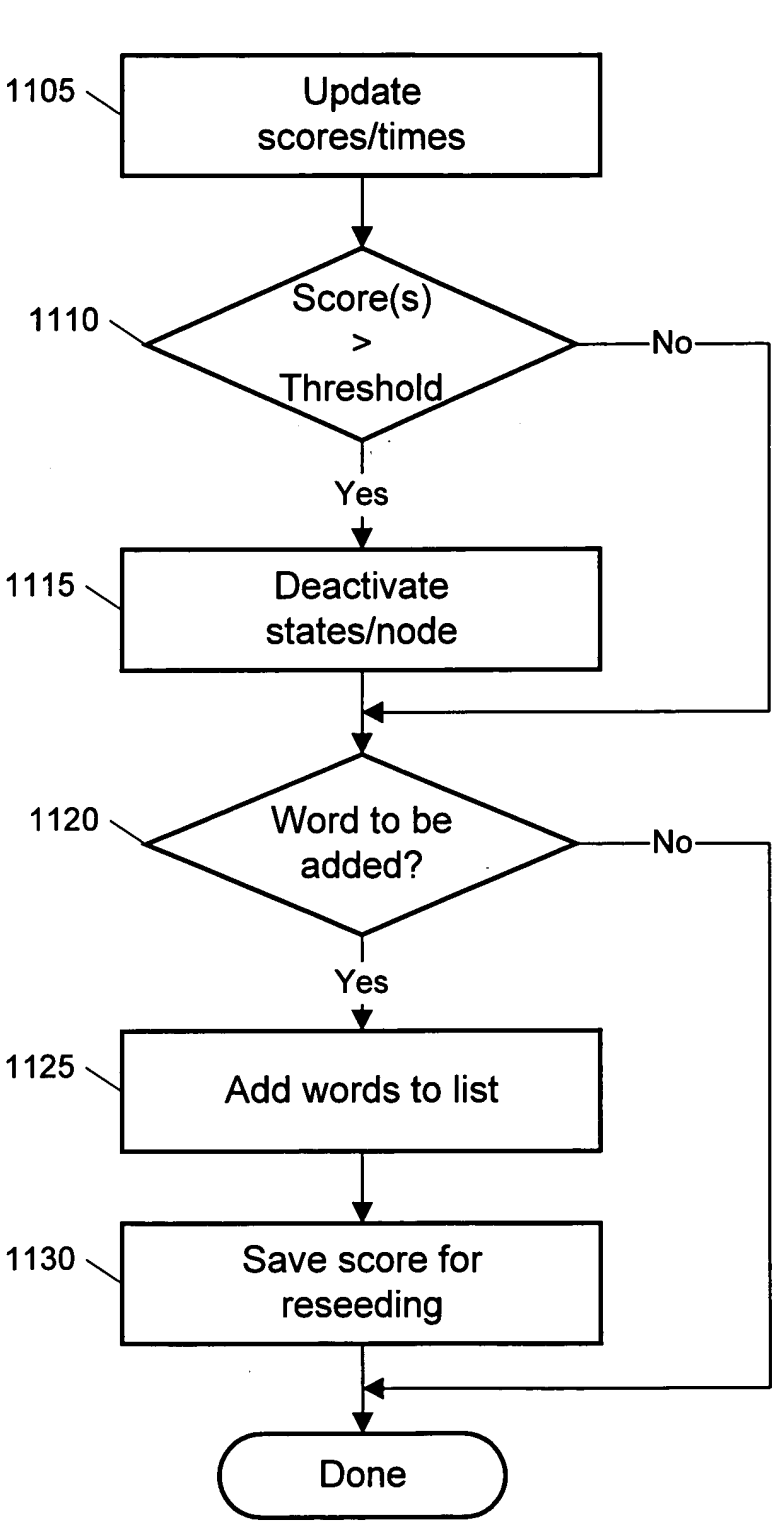
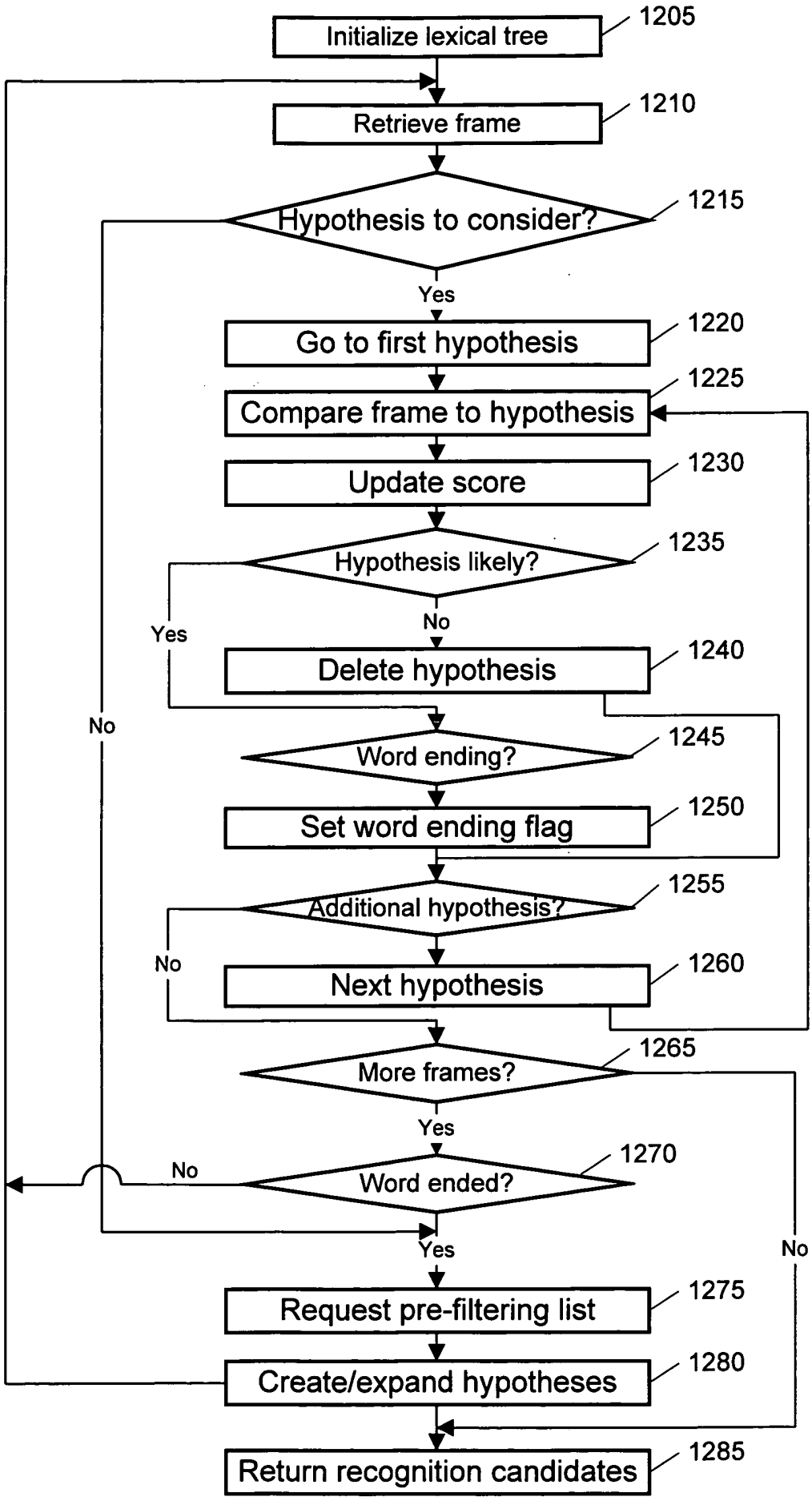

1100
1105
Update scores/times
1110
Score(s) > Threshold
No
Yes
1115
Deactivate states/node
1120
Word to be added?
No
Yes
1125
Add words to list
1130
Save score for reseeding
Done

Fig. 11

Fig. 12


Initialize lexical tree
1205
Retrieve frame
1210
Hypothesis to consider?
1215
Yes
No
Go to first hypothesis
1220
Compare frame to hypothesis
1225
Update score
1230
Hypothesis likely?
1235
Yes
No
Delete hypothesis
1240
Word ending?
1245
Set word ending flag
1250
Additional hypothesis?
1255
No
Next hypothesis
1260
More frames?
1265
Yes
No
Word ended?
1270
No
Yes
Request pre-filtering list
1275
Create/expand hypotheses
1280
Return recognition candidates
1285

1300 — "When a justice needs a friend New-Paragraph"

1315 — "there are two kinds of legal kibitzers"

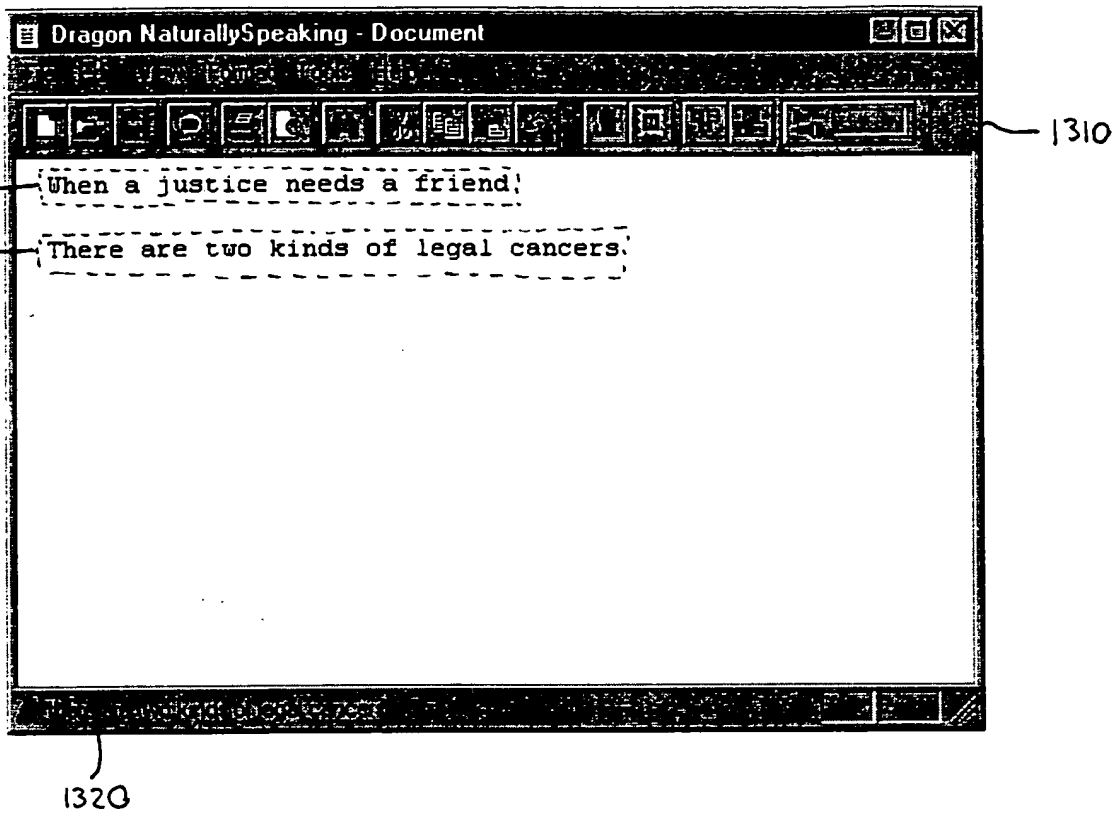


Fig. 13A

select cancers. "Spell That k i b i"

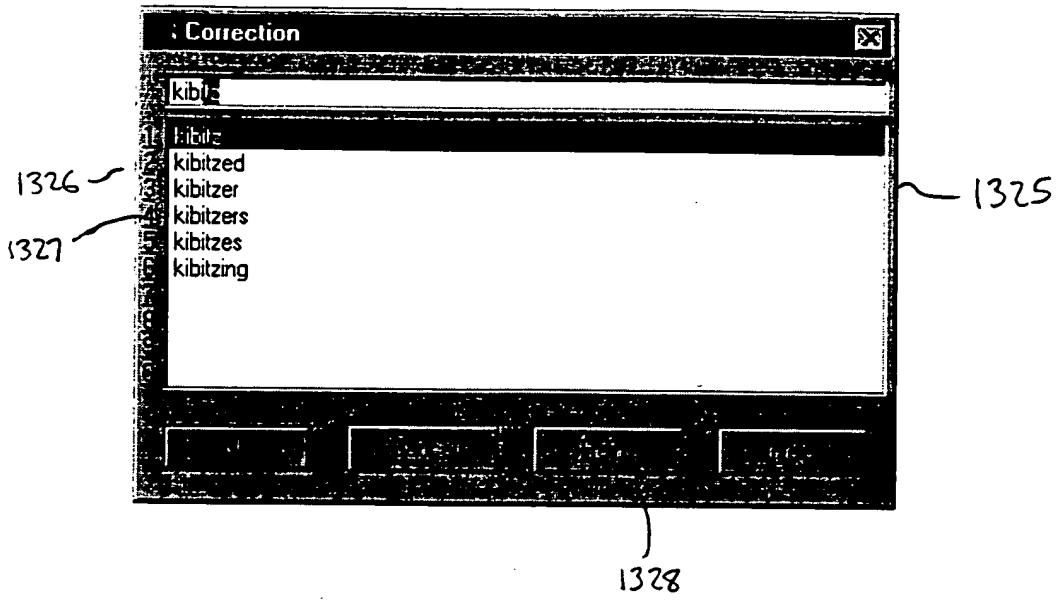


Fig. 13B

"Choose 4"

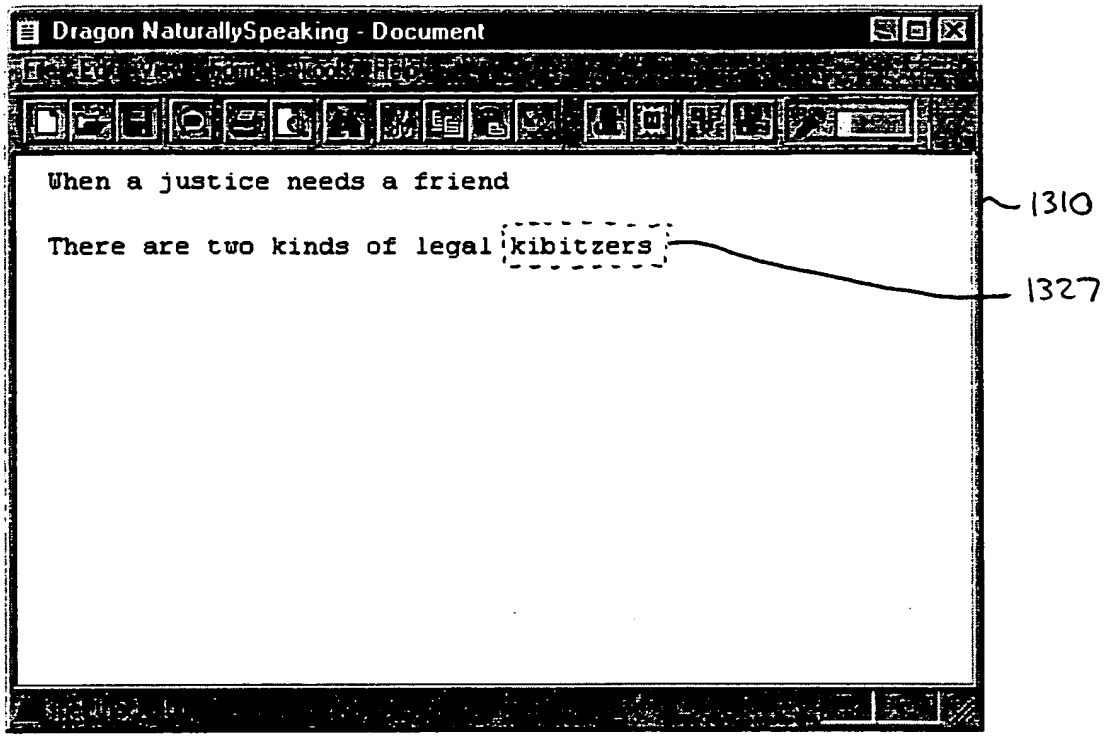


Fig. 13C

1329 ~ "those who pronounce amicus"

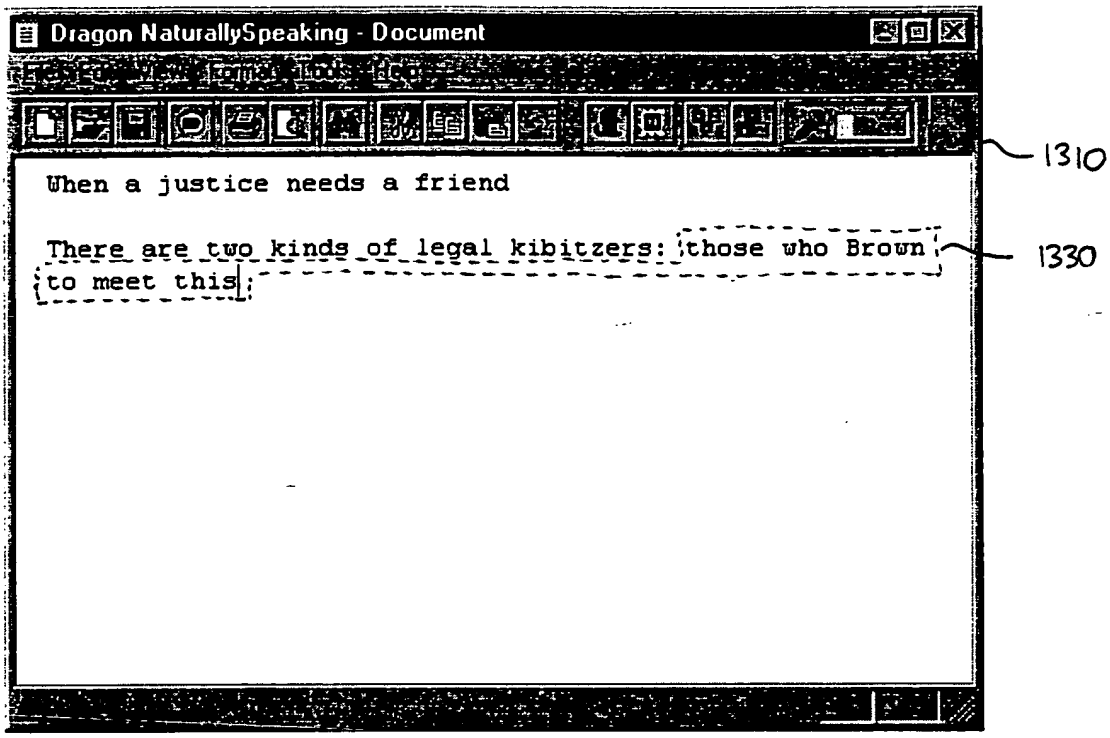


Fig. 13D

1331 — "Correct That"

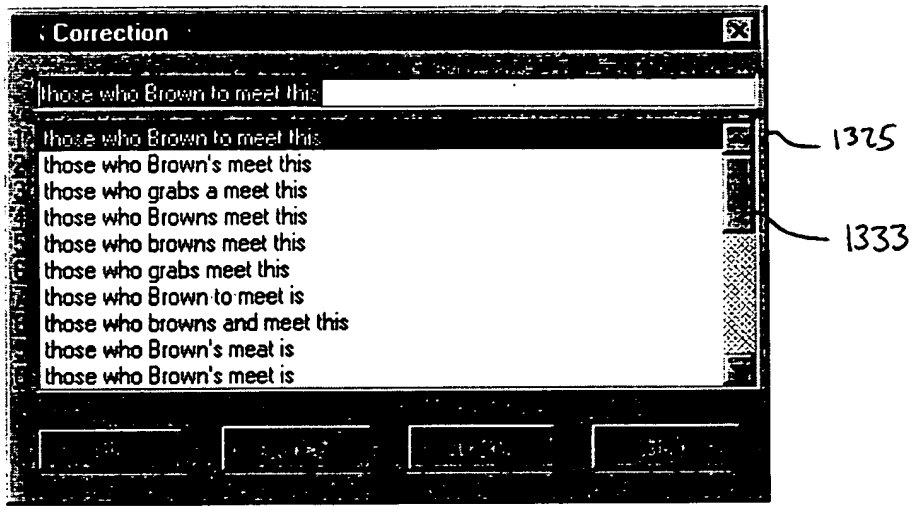


Fig. 13E

select "Brown" by mouse

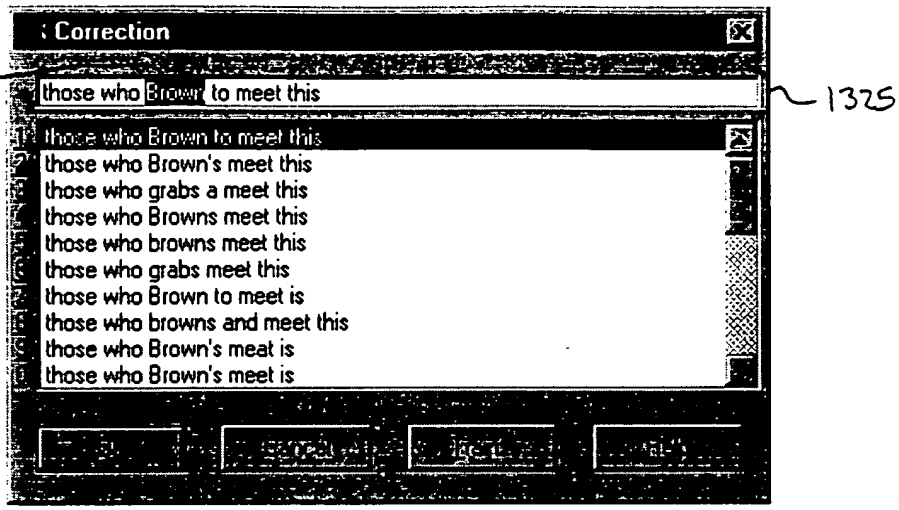


Fig. 13F

Docket No. 10559-0003

1340 "p r o n"

-- note that every choice ends with "to meet this".

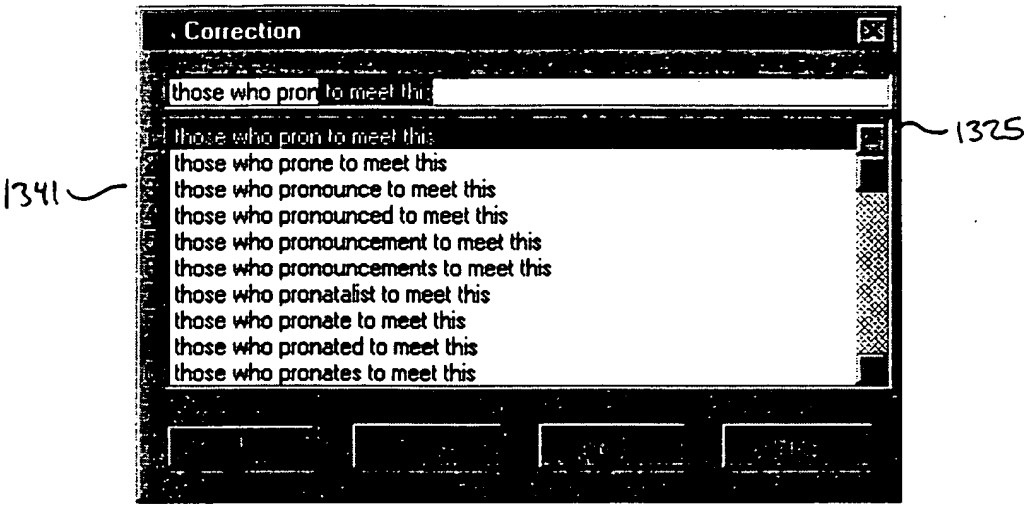


Fig. 136

Click on #3:

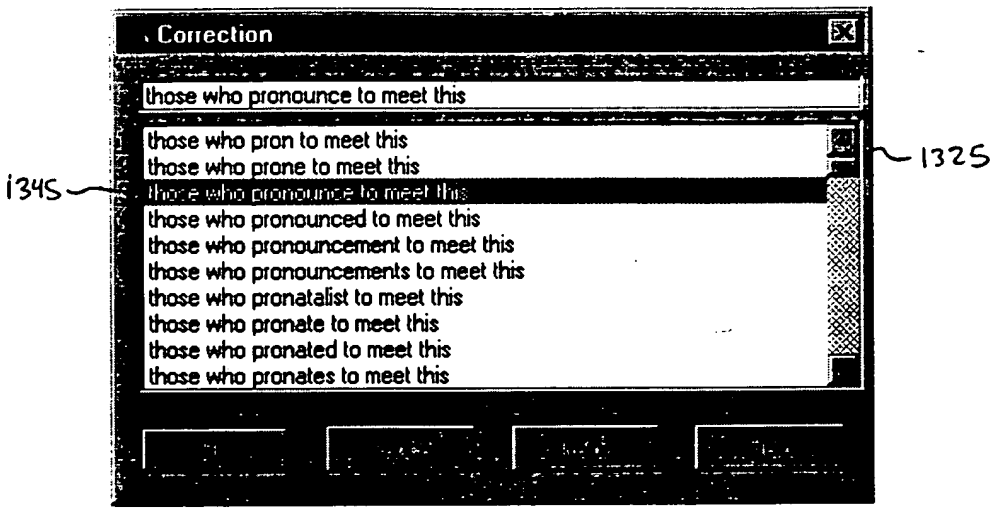


Fig. 13H

select "to meet this" by mouse

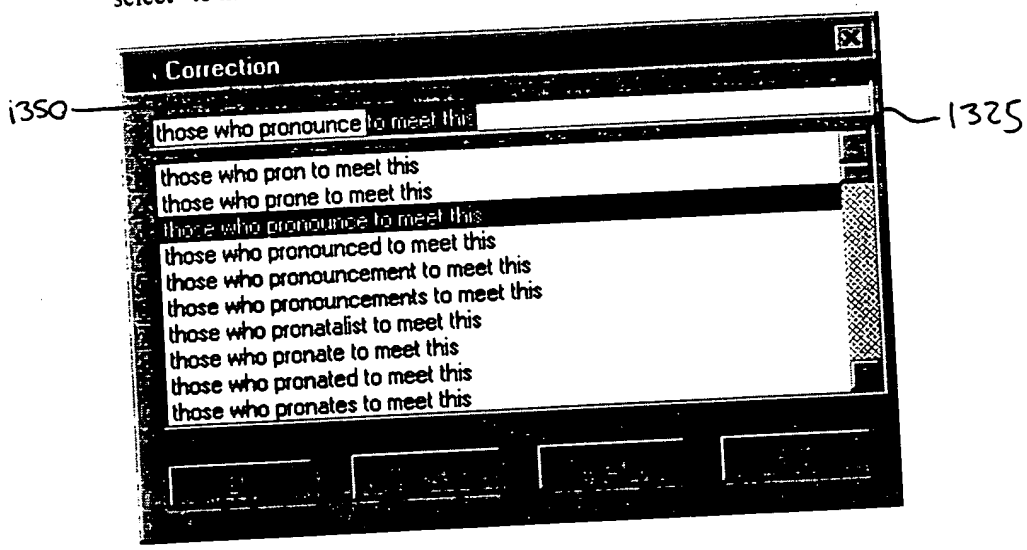


Fig. 13I

type "amicu"

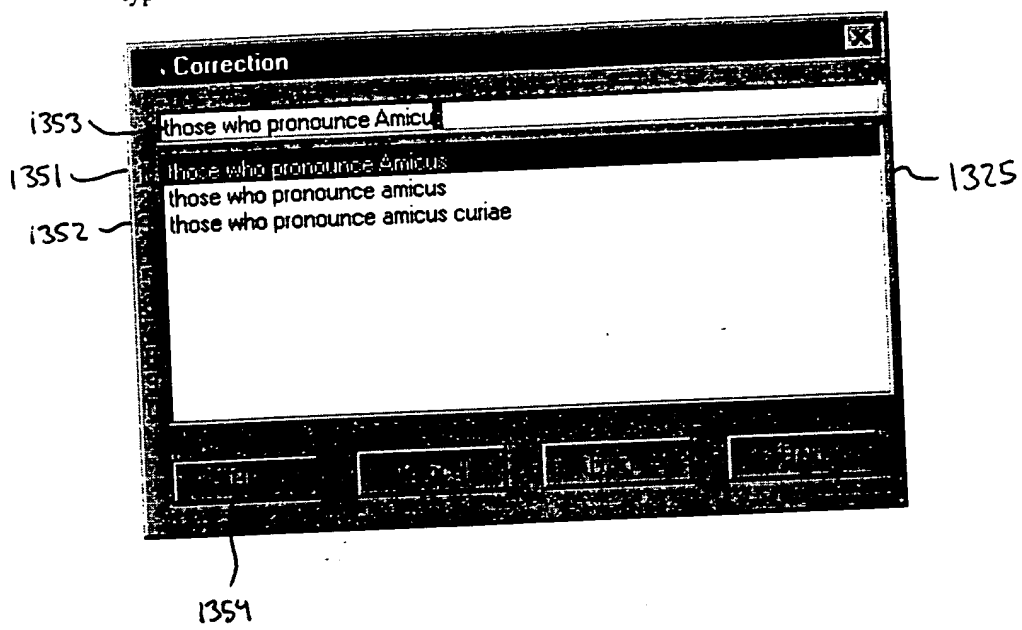


Fig. 13J

Click on OK.

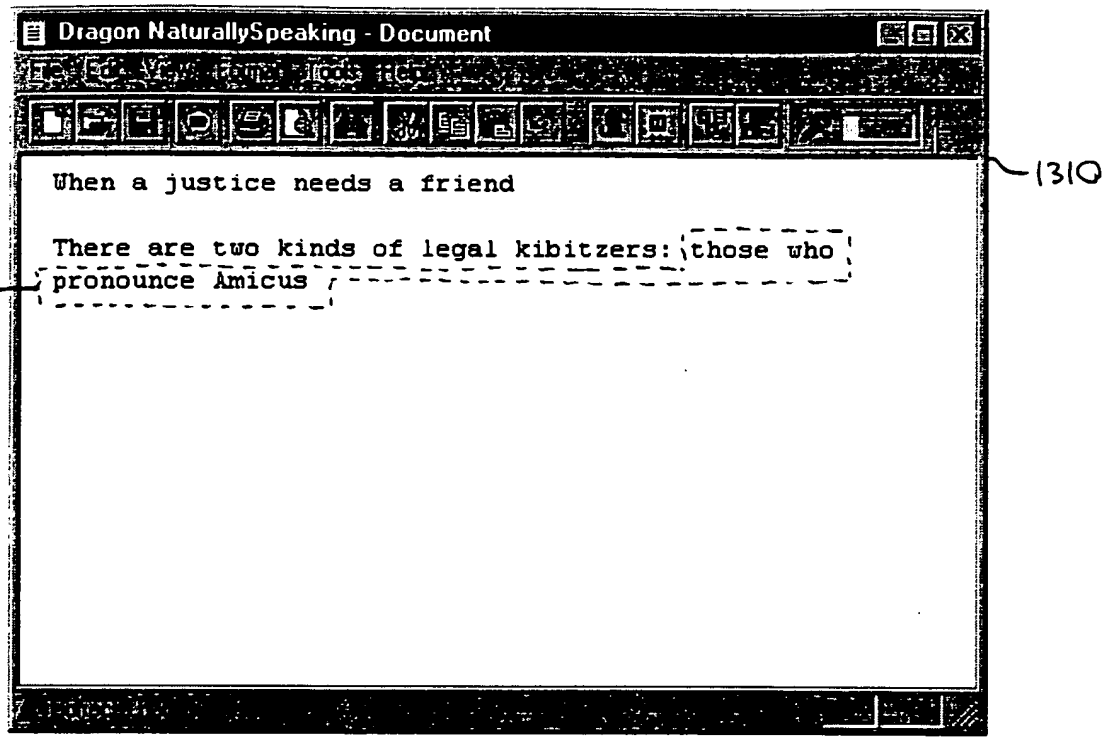


Fig. 13K

1360 — "Each submits a brief as an outsider"

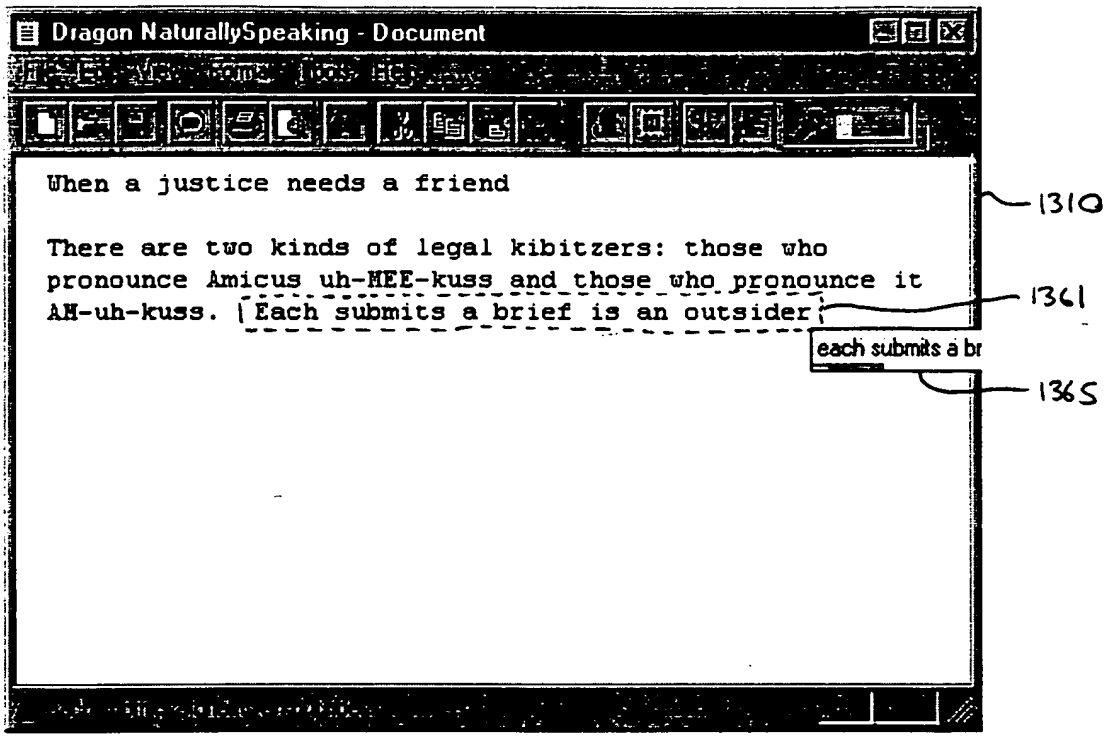


Fig. 13C

1331 — "Correct That"

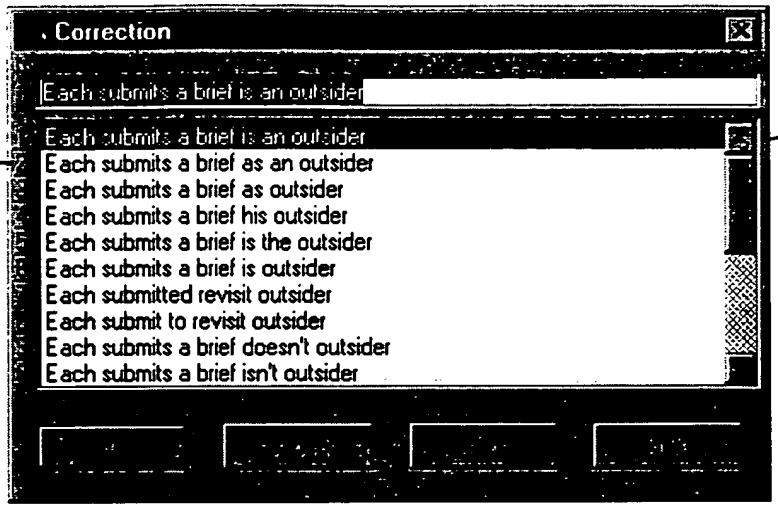


Fig. 13M

"Choose 2"

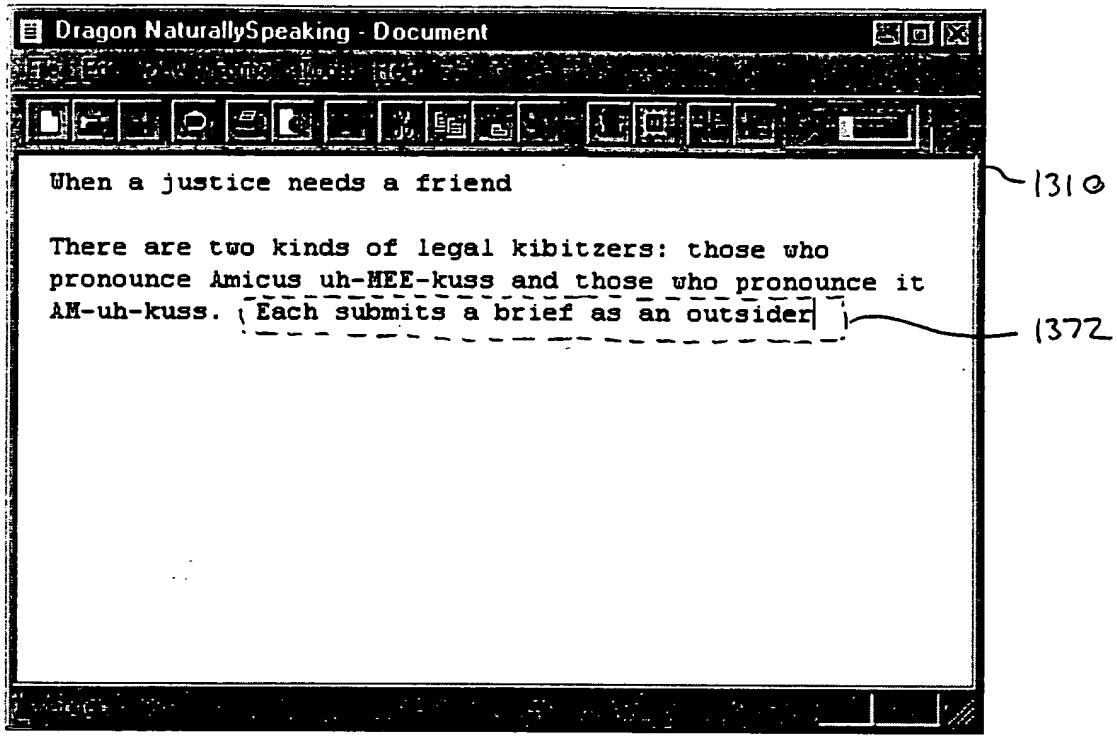


Fig. 13N

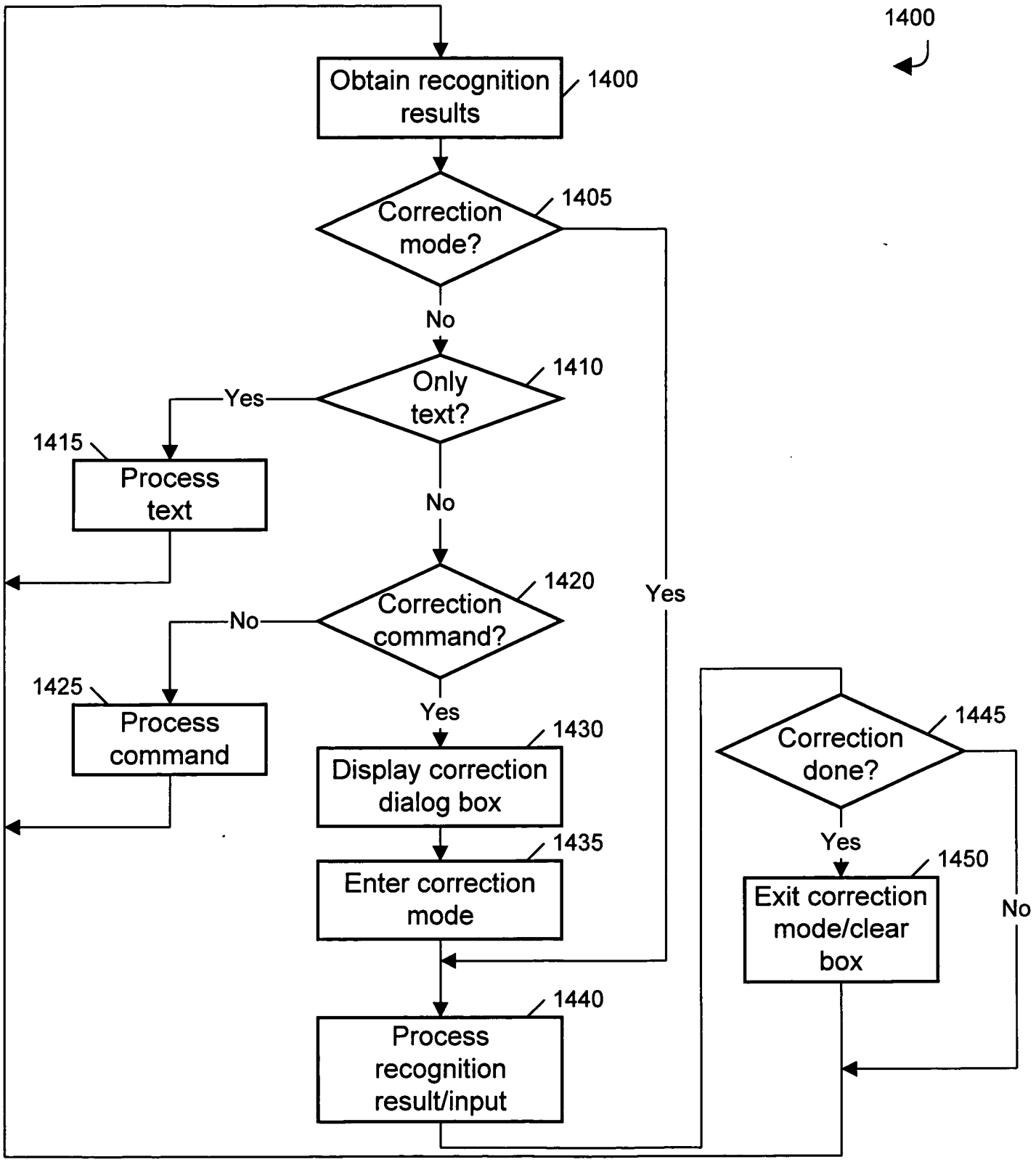

1400
Obtain recognition results
1400
Correction mode?
1405
No
Yes
Only text?
1410
1415
Process text
No
Correction command?
1420
No
Yes
1425
Process command
Yes
Display correction dialog box
1430
Enter correction mode
1435
Process recognition result/input
1440
Yes
Correction done?
1445
Yes
Exit correction mode/clear box
1450
No

Fig. 14

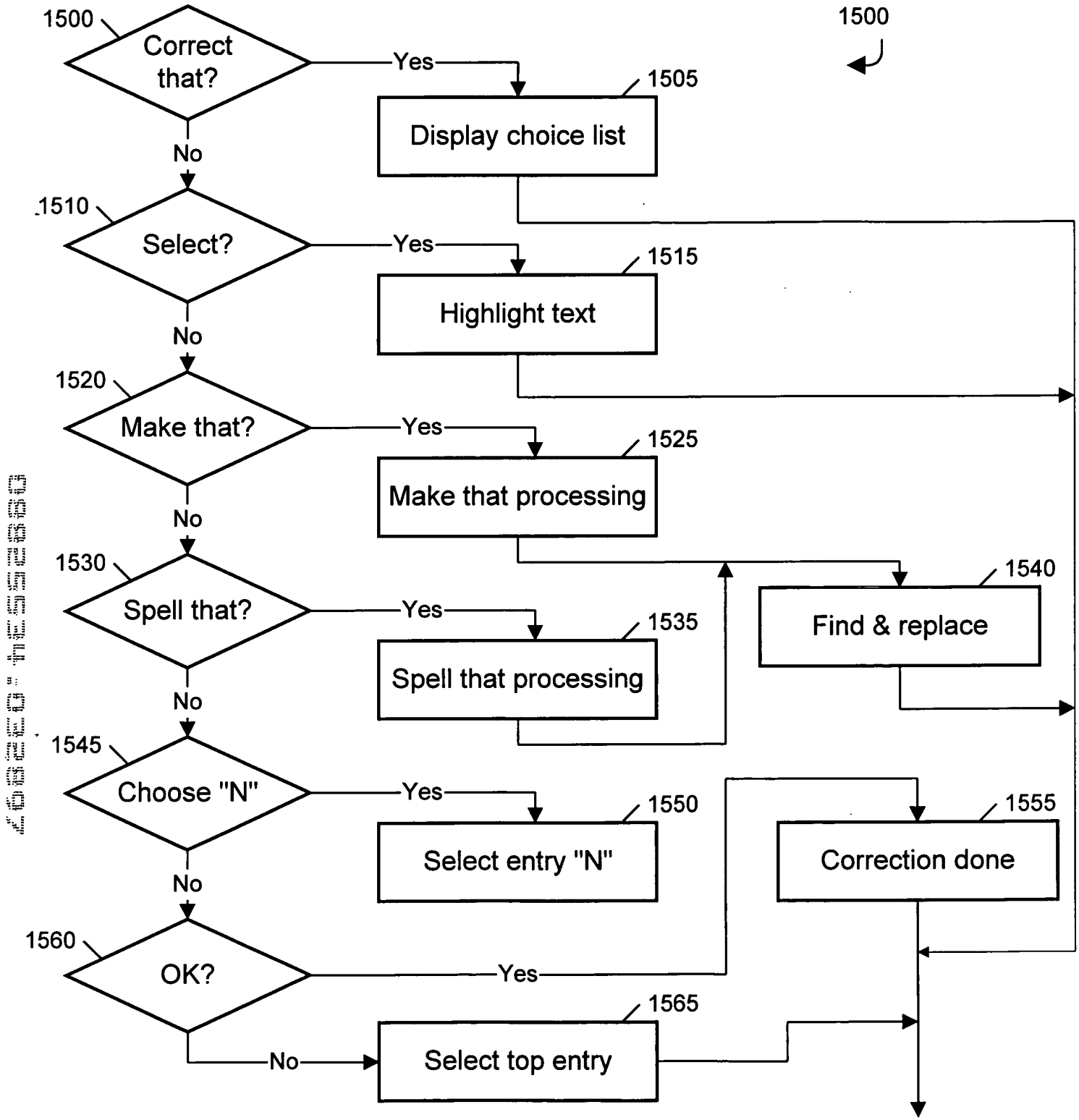

1500
Correct that?
Yes
1505
Display choice list
No
1510
Select?
Yes
1515
Highlight text
No
1520
Make that?
Yes
1525
Make that processing
No
1530
Spell that?
Yes
1535
Spell that processing
1540
Find & replace
No
1545
Choose "N"
Yes
1550
Select entry "N"
1555
Correction done
No
1560
OK?
Yes
1565
Select top entry
No

Fig. 15

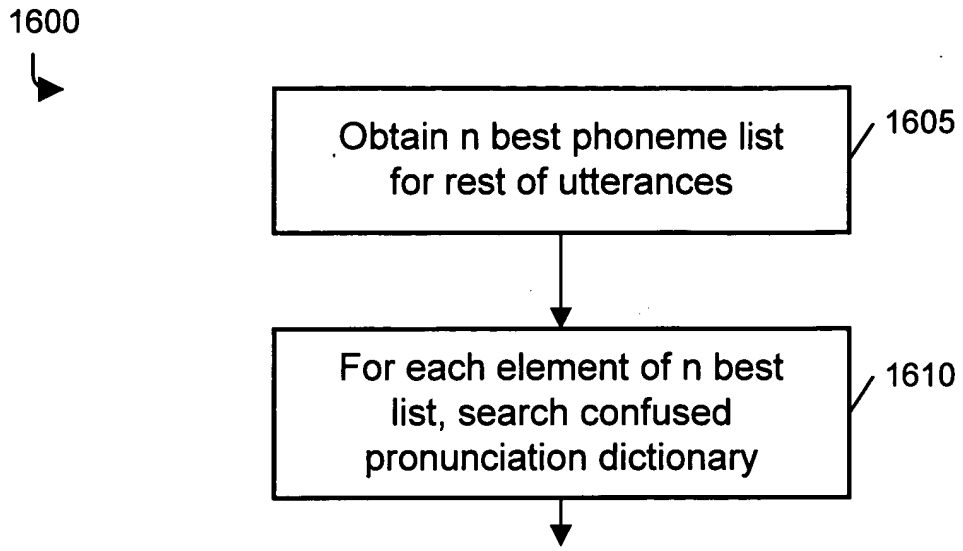


Fig. 16 (MAKE THAT)

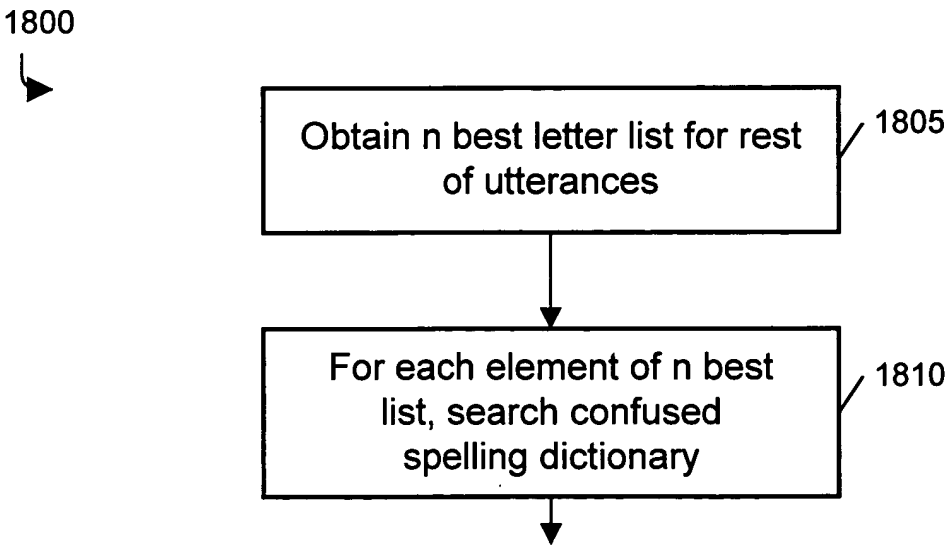


Fig. 18 (SPELL THAT)

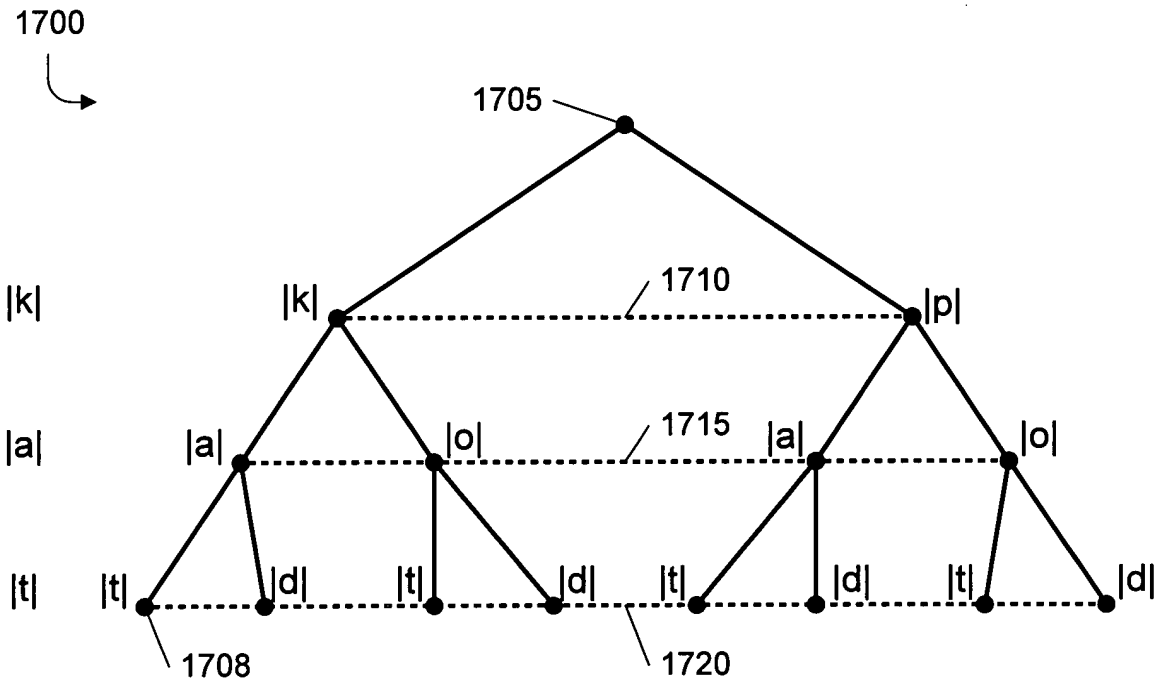


Fig. 17

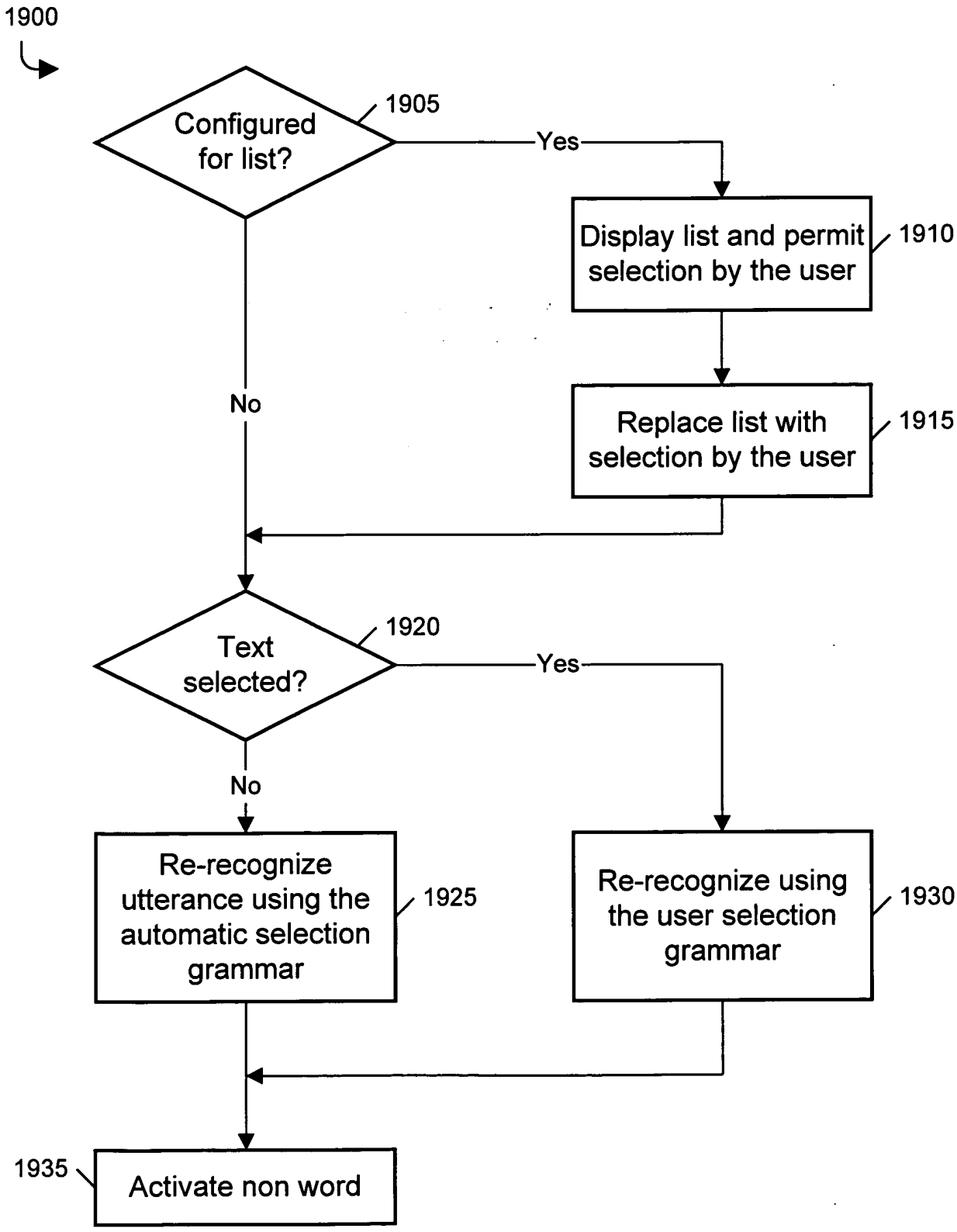

1900
Configured for list?
1905
Yes
No
Display list and permit selection by the user
1910
Replace list with selection by the user
1915
Text selected?
1920
Yes
No
Re-recognize utterance using the automatic selection grammar
1925
Re-recognize using the user selection grammar
1930
1935
Activate non word

Fig. 19

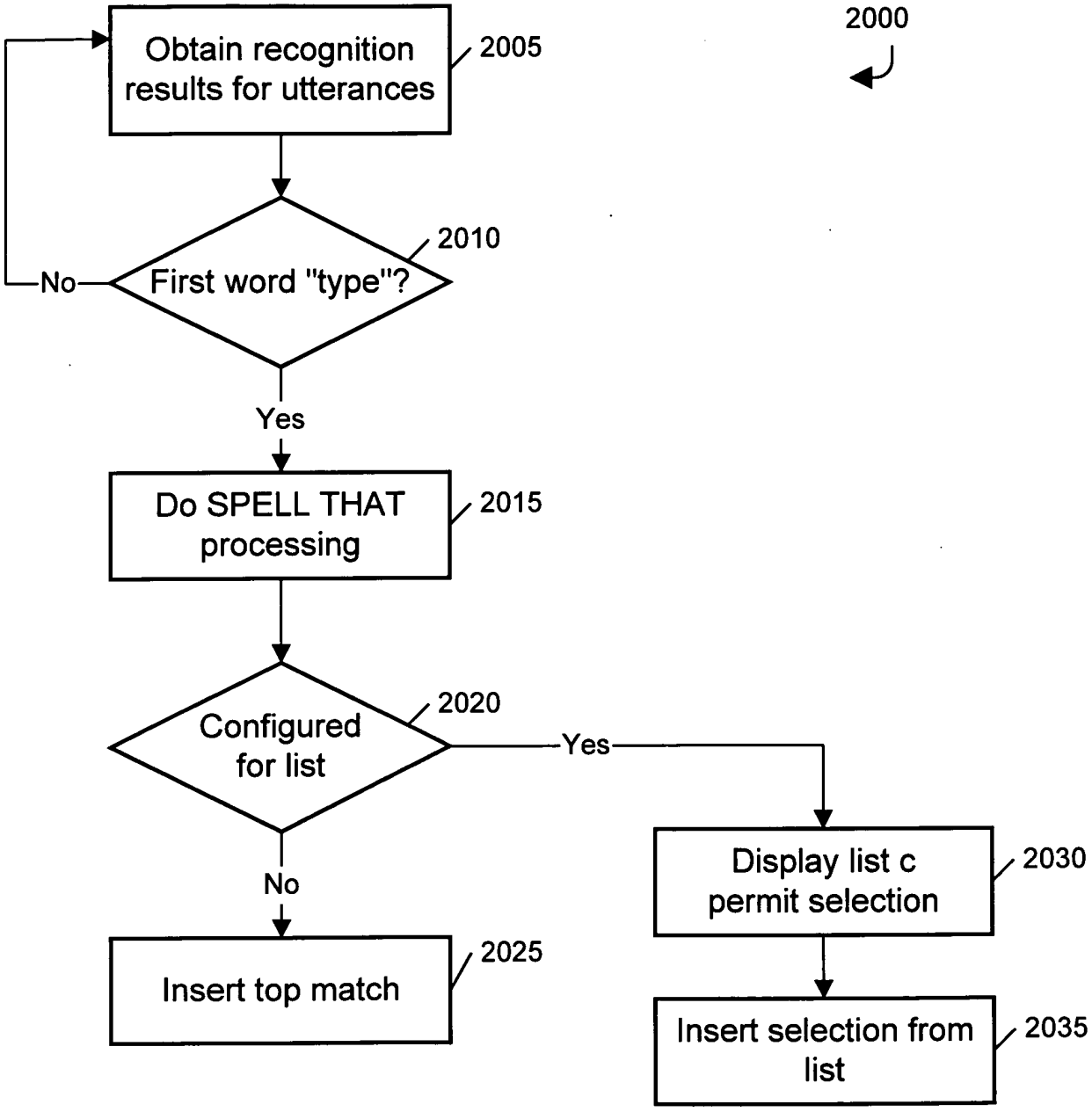

2000
Obtain recognition results for utterances
2005
First word "type"?
2010
No
Yes
Do SPELL THAT processing
2015
Configured for list
2020
Yes
No
Insert top match
2025
Display list c permit selection
2030
Insert selection from list
2035

Fig. 20 (TYPE)